

广义 Rough 近似关系基本性质及其计算^{*}

叶小平, 汤庸, 周风华, 蒋贻铨

(中山大学计算机科学系, 广东 广州 510275)

摘 要: 讨论了基于相似关系的广义 Rough 近似关系。通过提出一系列基本概念, 研究了广义 Rough 近似关系的“自反性”、“对称性”、“传递性”和“广义反对称性”等基本性质, 同时探讨了广义 Rough 关系的相关计算问题。

关键词: 相似关系; 广义 Rough 近似关系; 基本性质和计算

中图分类号: TP311.131 **文献标识码:** A **文章编号:** 0529-6579 (2004) 03-0001-04

关系是一种重要的集合类型, 一般集合的 Rough 集理论与方法^[1]自然也适用于关系的相关讨论。无论是从理论研究角度还是从实际应用观点出发, 人们更为关注的是关系所具有的不同于一般集合的特征与性质, 这在 Rough 集合中也是如此。在 Rough 关系研究中, 存在两个层面上的关系。一是近似空间 $AS = (U, \tilde{n})$ 中的非空个体域 U , 这是个笛卡尔乘积, 研究的对象 R 作为 U 的子集是一个关系; 一是个体域 U 上的关系 $\tilde{n}$, 这是一个与 U 处于不同层面的关系。Rough 关系理论就是基于近似空间 $(U, \tilde{n})$, 通过 U 上的关系 $\tilde{n}$ 展开对 U 中子集 R ——也就是关系 R 的研究。按照近似空间 $(U, \tilde{n})$ 中关系 $\tilde{n}$ 为等价关系或非等价关系可以将关系 R 的讨论分为传统 Rough 关系理论与广义 Rough 关系理论。传统 Rough 关系人们多有探讨^[2]。在各种重要的问题当中, 非等价关系情形并非少见, 例如不完备信息系统和基于变精度 Rough 理论的决策系统就是如此^[3]。讨论广义 Rough 关系具有理论与应用的必要性。Rough 理论的基本出发点是用关系 $\tilde{n}$ 将 U 中的元素进行分类。如果 $\tilde{n}$ 不具有自反性, 可以出现 U 中的元素 (x_1, x_2) 不在 $\tilde{n}$ 自身关系类 $\tilde{n}((x_1, x_2))$ 之中的情况, 进而可以有 (x_1, x_2) 不在 $\tilde{n}$ 的任何一个关系类之中。如果 $\tilde{n}$ 没有对称性, 就需要区分当 U 中的两个元素 (x_1, x_2) 和 (y_1, y_2) 有关系 $\tilde{n}$ 时, 是以 (x_1, x_2) 为基准的“前向”关系 $(x_1, x_2) \tilde{n} (y_1, y_2)$, 还是具有以 (x_1, x_2) 为基准的“后向”关系 $(y_1, y_2) \tilde{n} (x_1, x_2)$ 。以上两种情形过于一般, 在理论研究和实际应用经常考虑的是非等价关系的一种较强形式,

即满足自反性和对称性的“相似 (相容)”关系。相似关系不具有传递性, 同一个关系类之中的两个元素可以没有“关系”, 进而一个元素能够属于两个不同的关系类, 这是相似关系与等价关系的重要差异。本文较为详细地讨论了在 $\tilde{n}$ 为相似关系时, 基于近似空间 $AS = (U, \tilde{n})$ 广义 Rough 关系中一些基本问题, 其中的主要工作是给出了若干必须的基础性概念, 通过这些概念研究了广义上、下近似关系的基本性质。其中, 提出广义反对称性的概念, 将经典关系的偏序性质推广到广义的 Rough 关系, 而偏序性质与等价性质一样, 在关系理论与应用中有着重要意义。另外, 针对相似关系的特征, 讨论了广义 Rough 关系的计算思路和必要的“除噪”与“去冗”方法。考虑到 Rough 集理论与方法早已成为研究与应用的热点, 本文假设读者熟悉 Rough 集理论的相关文献, 例如文献 [1]。

1 基本概念

定义 1 设 n 是 U 上的关系。定义 $U = U \times U$ 上的关系 $\tilde{n} = n \times n$ 为

$$(x_1, x_2) \tilde{n} (y_1, y_2) \Leftrightarrow x_1 n y_1 \wedge x_2 n y_2$$

设 n 是 U 相似关系, 由定义 1 即知 $\tilde{n}$ 也是 $U = U \times U$ 上的相似关系。此时, 记 $\tilde{n}((x_1, x_2))$ 为 $(x_1, x_2) \in U \times U$ 所在的相似关系类, 可以证明 $\tilde{n}((x_1, x_2)) = n(x_1) \times n(x_2)$ 。以下常以 $n(x_1) \times n(x_2)$ 表示 (x_1, x_2) 在 $\tilde{n}$ 下的相似关系类 $\tilde{n}((x_1, x_2))$, 这里, $n(x)$ 是 x 在 n 之下关系类。

定义 2 设 U 是非空个体域, n 是 U 上的关系,

* 收稿日期: 2003-07-01

基金项目: 国家自然科学基金资助项目 (10171112)

作者简介: 叶小平 (1955 年生), 男, 副教授; E-mail: mcsxyp@zsu.edu.cn

$U = U \times U, \tilde{n} = n \times n$ 则称二元组: $AS = (U, \tilde{n}) = (U \times U, n \times n)$ 为一个广义近似空间。

定义 3 设 $AS = (U, \tilde{n})$ 是一个广义近似空间, R 是 U 的一个子集, 即 $R \subseteq U \times U$, 分别定义 R 的广义下近似关系 $L(AS, R)$ 和广义上近似关系 $H(AS, R)$ 如下:

$$L(AS, R) =$$

$$\{(x_1, x_2) \in U \times U; n(x_1) \times n(x_2) \subseteq R\}$$

$$H(AS, R) =$$

$$\{(x_1, x_2) \in U \times U; n(x_1) \times n(x_2) \cap R \neq \emptyset\}$$

广义下、上近似关系简称为下、上近似关系。

定义 4 设 $L(AS, R)$ 和 $H(AS, R)$ 为 $R \subseteq U \times U$ 的下、上近似关系, 称关系 $BND_{\tilde{n}}(R) = L(AS, R) - H(AS, R)$ 为 R 的边界关系。若 $BND_{\tilde{n}}(R) = \emptyset$, 则称 R 是广义 $\tilde{n}$ -可定义关系, 否则称 R 为广义 $\tilde{n}$ -Rough 关系, 简称为广义 Rough 关系。

定义 5 设 I 是 $U \times U$ 的子集。若 I 满足 $I = \bigcup_{x \in U} \tilde{n}((x, x))$, 则称 I 为基于关系 $\tilde{n}$ 的广义对角线关系, 简称广义对角关系。若以 R 记(常义) 对角线关系, 即 $R = \{(x, x) \in U \times U; \forall x \in U\}$, 则 $R \subseteq I$ 。

2 广义 Rough 近似关系基本性质

命题 1

(1) 设 R 是 $U \times U$ 的一个子集, 则 $L(AS, R)$ 具有自反性质的充要条件是 $I \subseteq R$ 。

(2) 若 R 具有自反性质, 则 $H(AS, R)$ 也具有自反性质。

证明

(1) 设 $L(AS, R)$ 具有自反性质, 即对任意 $(x, x) \in U \times U$ 成立 $n(x) \times n(x) \subseteq L(AS, R)$; 而此时 $L(AS, R) \subseteq R$, 所以 $I = \bigcup_{x \in U} \tilde{n}((x, x)) \subseteq R$, 必要性得证。反之, 如果 $I \subseteq R$, 则对任意 $(x, x) \in U \times U$ 成立 $n(x) \times n(x) \subseteq R$, 由 $L(AS, R)$ 的定义, 就有 $n(x) \times n(x) \subseteq L(AS, R)$, 所以 $L(AS, R)$ 具有自反性质。

(2) 注意到 $(x, x) \in n(x) \times n(x)$, 同时由题设 R 自反可知 $(x, x) \in R$, 这样就有 $n(x) \times n(x) \cap R \neq \emptyset$, 即 $(x, x) \in H(AS, R)$, 所以 $H(AS, R)$ 具有自反性质。

命题 2 若 R 具有对称性质, 则

(1) $L(AS, R)$ 具有对称性质。

(2) $H(AS, R)$ 具有对称性质。

证明

(1) 设 $(x_1, x_2) \in L(AS, R)$, 即 $n(x_1) \times n(x_2) \subseteq R$ 。

注意到任意 $(y_2, y_1) \in n(x_2) \times n(x_1) \Leftrightarrow (y_1, y_2) \in n(x_1) \times n(x_2)$, 由此可知若 $(y_1, y_2) \in R$, 由题设 R 具有对称性质, 故有 $(y_2, y_1) \in R$, 即 $n(x_2) \times n(x_1) \subseteq R$, 亦即 $(x_2, x_1) \in L(AS, R)$, 所以 $L(AS, R)$ 具有对称性质。

(2) 设 $n(x_1) \times n(x_2) \cap R \neq \emptyset$ 即存在 $(y_1, y_2) \in n(x_1) \times n(x_2) \cap R$ 。由 $(y_1, y_2) \in R$, 而 R 具有对称性质, 可得 $(y_2, y_1) \in R$, 即 $(y_2, y_1) \in n(x_2) \times n(x_1) \cap R$, 这样 $n(x_2) \times n(x_1) \cap R \neq \emptyset$ 。所以 $(x_2, x_1) \in H(AS, R)$, 即 $H(AS, R)$ 具有对称性质。

由命题 1 和命题 2 可已得到下面的定理 1。

定理 1

(1) 如果 $I = \bigcup_{x \in U} \tilde{n}((x, x)) \subseteq R$ 且 R 具有对称性质, 则 $L(AS, R)$ 是相似关系。

(2) 如果 R 是相似关系, 则 $H(AS, R)$ 也是相似关系。

命题 3 若 R 具传递性质, 则

(1) $L(AS, R)$ 具传递性质。

(2) 如 $I \subseteq R$, $H(AS, R)$ 具传递性。

证明

(1) 若 $(x, y) \in L(AS, R) \wedge (y, z) \in L(AS, R)$, 即 $n(x) \times n(y) \subseteq R \wedge n(y) \times n(z) \subseteq R$, 此时 $\forall (x', z') \in n(x) \times n(z)$, 取 $y' \in n(y)$, 就有

$$(x', y') \in n(x) \times n(y) \subseteq R \wedge$$

$$(y', z') \in n(y) \times n(z) \subseteq R$$

由 R 的传递性可知 $(x', z') \in R$, 所以, $n(x) \times n(z) \subseteq R$, 即 $(x, z) \in L(AS, R)$ 。

(2) 若 $(x, y) \in H(AS, R) \wedge (y, z) \in H(AS, R)$ 即 $n(x) \times n(y) \subseteq R \neq \emptyset \wedge n(y) \times n(z) \cap R \neq \emptyset$ 。

取 $(x', y') \in n(x) \times n(y) \cap R$ 且 $(y'', z') \in n(y) \times n(z) \cap R$ 。由 $y' \in n(y) \wedge y'' \in n(y)$ 可知, $(y', y'') \in n(y) \times n(y) \subseteq I$; 由题设 $I \subseteq R$, 则 $(y', y'') \in R$ 。这样就有 $(x', y') \in R$ 且 $(y'', z') \in R$ 且 $(y', y'') \in R$ 。由题设 R 具传递性质, 故 $(x', z') \in R$, 即 $(x', z') \in n(x) \times n(z) \cap R$, 从而 $n(x) \times n(z) \cap R \neq \emptyset$, 故有 $(x, z) \in H(AS, R)$ 。

命题 3 说明, 广义近似关系的传递性质与广义近似空间 $AS = (U, \tilde{n})$ 中相似关系 $\tilde{n}$ 是否具有传递性质是无关的, 这是一个有趣的结果。由命题 1、命题 2 和命题 3, 可以得到下面定理 2。

定理 2

(1) 如果 $I \subseteq R$ 且 R 具有对称性质和传递性质, 则 $L(AS, R)$ 是等价关系。

(2) 如果 R 是等价关系, 并且 $I \subseteq R$, 则 $H(AS,$

R) 是等价关系。

定义 6 设 $R \subseteq U \times U$, 若对 $(x_1, x_2) \in R$, 当 $(x_2, x_1) \in R$ 时, 必有 $(x_1, x_2) \in I$, 则称 R 具有在 $\bar{n}$ 意义下的反对称性质, 简称 R 具有广义反对称性质。

广义反对称性质是通常反对称性质的推广, 因为若 R 通常反对称关系, 即当 $(x_1, x_2) \in R$ 并且 $(x_2, x_1) \in R$, 必有 $x_1 = x_2$, 此时也有 $(x_1, x_2) \in I$, 所以 R 具广义反对称性, 但反之不然。

如果 R 具有自反性, 广义反对称性和传递性, 则称 R 是基于近似空间 AS 的偏序关系, 简称 R 是广义偏序关系。

命题 4 设 R 具有反对称性质, 则 (1) $L(AS, R)$ 也具有反对称性质; (2) 若 $H(AS, R) \subseteq R \cup I$, 则 $H(AS, R)$ 具有广义反对称性质。

证明

(1) 设 $(x_1, x_2) \in L(AS, R)$ 且 $(x_2, x_1) \in L(AS, R)$, 即 $n(x_1) \times n(x_2) \subseteq R$ 且 $n(x_2) \times n(x_1) \subseteq R$, 则 $(x_1, x_2) \in R$ 且 $(x_2, x_1) \in R$ 。由 R 具反对称性可知 $x_1 = x_2$, 所以 $L(AS, R)$ 也具反对称性。

(2) 设 $(x_1, x_2) \in H(AS, R) \subseteq R \cup I$ 且 $(x_1, x_2) \notin I$, 此时 $(x_1, x_2) \in R$ 。注意由 $(x_1, x_2) \notin I$ 可得 $x_1 \neq x_2$ 。由 R 反对称性 $(x_2, x_1) \notin R$, I 是相似关系, 也有 $(x_2, x_1) \notin I$ (否则就有 $(x_1, x_2) \in I$, 出现矛盾), 所以 $(x_2, x_1) \notin H(AS, R)$ 。由定义 6, $H(AS, R)$ 具广义反对称性质。

说明 一般而言, R 具有反对称性质, 不能保证 $H(AS, R)$ 具有反对称性质, 而命题 4 给出了 $H(AS, R)$ 具有广义反对称性质的一个较为宽泛的充分条件。下面命题讨论了广义反对称的性质。

命题 5 若 $H(AS, R)$ 具有广义反对称性质, 则 $I \cap R \neq \phi$ 。

证明 设 $(x_1, x_2) \in H(AS, R)$, 且 $(x_2, x_1) \in H(AS, R)$, 即 $n(x_1) \times n(x_2) \cap R \neq \phi$ 和 $n(x_2) \times n(x_1) \cap R \neq \phi$ 。由 $H(AS, R)$ 具有广义反对称性质可得 $(x_1, x_2) \in I$, 所以 $I \cap R \supseteq \{(x_1, x_2)\}$, 即 $I \cap R \neq \phi$ 。

由命题 1、命题 3 和命题 4, 可以得到下面的定理 3。

定理 3 如果 $I \subseteq R$ 且 R 具有反对称性质和传递性质, 则 $L(AS, R)$ 是偏序关系。

3 上、下广义近似关系的计算

在 Rough 集理论中, 一个信息系统就是一个近似空间。其中的属性就是相应近似空间的关系^[2,3]

显然, 有意义的信息系统中可以有不止一个属性, 对应的关系一般应当是一个关系族。传统 Rough 关系需要由等价关系族产生一个新的等价关系——不分明关系, 进而得到所讨论集合的下、上近似关系。在广义 Rough 集理论中也需要考虑类似的问题:

① 如何由相似关系族得到产生“上、下近似集”所必需的“广义不分明关系”;

② 如何对相应的上、下近似集进行必要的“去噪”和“除冗”。由于广义不分明关系是由相似关系族产生, 一般而言, 仍然是一个相似关系。相似关系不具备传递性, 得到的关系类仅是论域的“覆盖”而非“划分”。由此产生的上、下近似集就具有所谓的“噪声”与“冗余”, 需要考虑此类问题的处理。

出于行文简洁明晰的考虑, 下面的算法仅对 Rough 集而言, 除了有关符号的不同之外, 对于 Rough 关系也是适用的。

定义 7 设 $(U, \mathbf{R})$ 为广义近似空间, $\mathbf{R}$ 为 U 上的相似关系族。对 $R \in \mathbf{R}, x, y \in U$, 若 xRy , 定义

$$\lambda_R(x, y) = \lambda(R(x), R(y)) = \frac{|R(x) \cap R(y)|}{\min(|R(x)|, |R(y)|)}$$

称其为 x, y 的 R 关系度, 其中 $|S|$ 表示集合“ S ”的基数。显然, $\lambda_R(x, y)$ 是定义域含于 $U \times U$ 的二元函数, 称其为 R 相似度函数。

如果 x 和 y “具有”相似关系 R , R 相似度函数“度量”了两个相似关系类 $R(x)$ 与 $R(y)$ 的“相似关联程度”, 是传统关系中“传递性”的定量刻画。由定义 7 可以知道下述命题成立。

命题 6 如果 R 具传递性质时, $\lambda_R(x, y) = 1$ 。

定义 8 设 $(U, \mathbf{R})$ 为广义近似空间, $\mathbf{R} = \{R\}$ 为 U 上的相似关系族, $\lambda_R(x, y)$ 是关系族 $\mathbf{R} = \{R\}$ 中 R 相似度函数列, 记 $f(x, y) = \sum_R \lambda_R(x, y)$, 并称其为关系族 $\mathbf{R}$ 的不分明关系函数。

定义 9 对每个 $\lambda_R(x, y)$, 取定阈值 $\hat{q}_R: 0 \leq \hat{q}_R \leq 1$ 。记 $\lambda = \sum_R \hat{q}_R$ 。在 U 上定义一个关系 R 如下:

$\forall x, y \in U, xRy$ 当且仅当 $f(x, y) \geq \lambda$ 。如此定义的关系 R 称为由关系族 $\mathbf{R} = \{R\}$ 产生的广义不分明关系。由这个定义可知下述命题成立。

命题 7

- (1) R 是 U 上的相似关系;
 - (2) 当 $\mathbf{R} = \{R\}$ 为等价关系族时, $R = \bigcap \mathbf{R}$ 。
- 为着简便, 下面讨论中总以 R 记 R 。

定义 10^[3] 设 R 是非空个体域 U 上的相似关

系, X 是 U 的子集. 对非负整数 k_1 和 k_2 , 定义

$$R_{\#}(X, k_1) =$$

$$\{x \in U: |R(x)| \vdash |R(x) \cap X| \leq k_1\}$$

$$R^{\#}(X, k_2) = \{x \in U: |R(x) \cap X| \geq k_2\}$$

分别称为 X 的 $k_1 - R$ 下近似集和 $k_2 - R$ 上近似集, 简称 K -下、上近似集. 由该定义可以得到:

命题 8^[3] 当 $k_1 = k_2 = 0$ 时, 得到传统意义下的下、上近似集.

由于相似关系 R 不具备传递性质, 如果直接使用传统上下近似集合, 可能使得下近似集“太小”, 而上近似集“过大”. 为此需要将下近似集适当“放大”, 上近似集适当“缩小”. 定义 10 实际上是放宽下近似集条件: 只要 $R(x)$ 中有不多于 k_1 个点不属于 X 就认为 $R(x)$ 被 X “包含”; 同时加强上近似集要求: 只有在 X 中的点多于 k_2 个的 $R(x)$ 才能认为是与 X 的“交”非空.

K 近似集可以看作具有某种消除干扰和除去冗余功能的基本集合.

如前所述, 对于相似关系 R 而言, 同一个对象可被多个“不同”的关系类重复“覆盖”, 应当除掉冗余, 其出发点是需保障除冗前后的“基本因

素”不变, 或者至少不能有难以容忍的改变. 基于这样的思路, 引入冗余关系类的概念.

定义 11 设 R 是 U 上的广义不可分辨关系, $x, y \in U, X \subseteq U, R_{\#}(X)$ 和 $R^{\#}(X)$ 分别是 X 的 K -下、上近似集, $R(x), R(y)$ 是两个相似关系类.

(1) 如果 $R(x), R(y) \subseteq R_{\#}(X)$, 且 $R(x) \cap R(y) \neq \phi$, 其中当 $|R(x)| \leq |R(y)|$, 则称 $R(x)$ 是关于 X 下冗余相似类.

(2) 如果 $R(x), R(y) \subseteq R^{\#}(X) - R_{\#}(X)$, 且 $R(x) \cap R(y) \neq \phi$, 当 $|R(x)| \leq |R(y)|$, 则称 $R(y)$ 是关于 X 上冗余相似类.

显然, 除去定义 11 中所描述的冗余关系类, X 的 K -下近似集不会改变, 而 K -上近似集也没有“大的”变动, 可以认为是合理的.

参考文献:

- [1] PAWLAK Z. Rough set[J]. International J Information and Computer Sci, 1982, 11(5): 341—356.
- [2] 刘清. Rough 集及 Rough 推理[M]. 北京: 科学出版社, 2001.
- [3] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001.

Properties and Calculation on Generalized Rough Approximate Relation

YE Xiao ping, TANG Yong, ZHOU Feng hua, JIANG Yi quan

(Department of Computer Science, Sun Yat sen University, Guangzhou 510275, China)

Abstract: The generalized Rough relation grounded on common similar relation is studied. After giving a serial of basic concepts, some important characters such as reflexive, symmetric, transfer, generalized partial order properties are studied.

Key words: similar relation; generalized approximate relation; basic properties and calculation